

*Sentient Structures*

**Optimising Sensor Layouts for Direct Measurement of Discrete Variables**

**Report to US Air Force Research Laboratory,  
Asian Office of Aerospace Research and Development (AOARD)  
AOARD Case Number AOARD-074045**



**CSIRO Materials Science and Engineering**

Report No. CMSE-2008-412

November 2008

Report Documentation Page			Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.				
1. REPORT DATE <b>02 DEC 2008</b>	2. REPORT TYPE <b>Final</b>	3. DATES COVERED <b>01-05-2007 to 01-05-2008</b>		
4. TITLE AND SUBTITLE <b>Sentient Structures</b>		5a. CONTRACT NUMBER <b>FA48690714045</b>		
		5b. GRANT NUMBER		
		5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) <b>Donald Price</b>		5d. PROJECT NUMBER		
		5e. TASK NUMBER		
		5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>CSIRO,PO Box 218,Lindfield, NSW ,Australia,NA,2070</b>		8. PERFORMING ORGANIZATION REPORT NUMBER <b>N/A</b>		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) <b>AOARD, UNIT 45002, APO, AP, 96337-5002</b>		10. SPONSOR/MONITOR'S ACRONYM(S) <b>AOARD</b>		
		11. SPONSOR/MONITOR'S REPORT NUMBER(S) <b>AOARD-084027</b>		
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT <b>Development of reliable methodologies for determination of optimal sensor placements is an important requirement for the development of sentient structures. An optimal sensor layout is attained when a limited number of sensors are placed in an area such that the cost of the placement is minimised while the value of the obtained information is maximised. In this report, we first introduce a criterion that maximizes the value, or expected benefit, of using a sensor subset for a given sensor model relative to the environment. Defining the value in terms of the information obtained allows the sensor layout problem to be represented as an entropy optimization problem. This criterion is compared with other well-known criteria, both theoretically and experimentally, the latter by comparing the various criteria for optimal sensor layout using data from an existing wireless sensor network. This is achieved by firstly learning a spatial model of the environment using a Bayesian Network architecture, then predicting the expected sensor data in the rest of the space, and lastly verifying the predicted results with actual measurements.</b>				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>Same as Report (SAR)</b>	18. NUMBER OF PAGES <b>35</b>
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>		

This Report represents the deliverable for AOARD contract no. FA4869-07-1-4045.

### CSIRO Staff Who Contributed to this Work

Rosalind Wang<sup>1</sup>  
George Mathews<sup>2,3</sup>  
Don Price<sup>2</sup>  
Mikhail Prokopenko<sup>1</sup>

Contact: Don Price  
Telephone: +61 2 9413-7447  
Email: [Don.Price@csiro.au](mailto:Don.Price@csiro.au)

### Abstract

Development of reliable methodologies for determination of optimal sensor placements is an important requirement for the development of sentient structures. An optimal sensor layout is attained when a limited number of sensors are placed in an area such that the cost of the placement is minimised while the value of the obtained information is maximised. In this report, we first introduce a criterion that maximises the value, or expected benefit, of using a sensor subset for a given sensor model relative to the environment. Defining the value in terms of the information obtained allows the sensor layout problem to be represented as an entropy optimisation problem. This criterion is compared with other well-known criteria, both theoretically and experimentally, the latter by comparing the various criteria for optimal sensor layout using data from an existing wireless sensor network. This is achieved by firstly learning a spatial model of the environment using a Bayesian Network architecture, then predicting the expected sensor data in the rest of the space, and lastly verifying the predicted results with actual measurements (ground truth).

---

<sup>1</sup> CSIRO Information & Communications Technology Centre

<sup>2</sup> CSIRO Materials Science & Engineering

<sup>3</sup> Australian Centre for Field Robotics, University of Sydney

## Table of Contents

<b>1. Introduction.....</b>	<b>4</b>
1.1 Direct and indirect sensing.....	4
1.2 Prior knowledge and environmental models.....	5
1.3 Sensor-environment scenarios .....	6
1.4 Cost and value.....	6
<b>2. Formulation of the Optimal Design Problem.....</b>	<b>7</b>
2.1 An illustrative scenario .....	8
2.2 Inclusion of sensors and observations.....	9
2.3 Optimal sensor layout .....	10
2.4 Information-theoretic measures .....	11
2.5 Graphical representations of information theoretic quantities.....	12
2.6 Optimal selection of a subset .....	13
<b>3. Current Methods.....</b>	<b>14</b>
3.1 Reward and entropy (RE) .....	14
3.2 Mutual information (MI).....	15
3.3 Information coverage (IC) .....	16
<b>4. Experimental Setup .....</b>	<b>16</b>
4.1 Data .....	16
4.2 Bayesian networks .....	18
4.3 Learning and inference .....	19
<b>5. Results and Discussion.....</b>	<b>20</b>
5.1 Two Sensors.....	20
5.2 Three and Four Sensors .....	23
<b>6. Discussion and Conclusions .....</b>	<b>24</b>
<b>References .....</b>	<b>26</b>

## 1. Introduction

The vision of engineering structures of the future having sentient properties was outlined in the White Paper submitted to AOARD prior to the commencement of this work, and which is attached as an appendix to this report. One of the essential requirements of a sentient structure is a distributed sensing capability that enables the structure to sense its own state, and properties of its environment that may affect its state or its required functions. Sensing is critical to the ability of the structure to be aware of its state and its surroundings, and consequently to its ability to respond.

Distributed sensing is becoming increasingly important in many areas of modern society, even where the vision of the future does not extend to sentient structures. Sensing for environmental monitoring (on scales from personal spaces – rooms, buildings, vehicles – to catchments, ecosystems, continents and oceans), industrial process monitoring, monitoring for security and safety, health and well-being, and structural health monitoring are examples of areas in which distributed sensing is becoming important.

Critical issues that must be addressed in the development of a distributed sensing system are what is to be sensed, where sensors should be located, and how should the data be communicated and processed. In many cases information is required on multiple scales, and a high density of sensors would be required to provide small-scale information over a large area or volume. In practice, limitations (sometimes severe) are placed on the number of sensors that may be deployed by considerations of cost, weight, sensor size, and the ability to communicate and process the large volume of data effectively. Therefore, it is important to find methods to place a limited number of sensor nodes in the area of interest such that the *cost* of the placement is minimised while the *value* of the obtained information is maximised.

As indicated in the White Paper (Appendix 1) an initial focus of the work of developing sentient structures will be the development of information-theoretic techniques for determining optimal sensor densities and layouts. This report describes an initial examination of this problem for the simple case of the direct measurement of discrete variables. There is much more to be done, as will be indicated in the final section of the report.

### 1.1 Direct and indirect sensing

The work outlined in this report considers only the simple case of sensors that make a direct, local measurement of a parameter of interest at the location of the sensor (e.g. a thermometer). Sensors that make non-local measurements (e.g. a radiation sensor if the quantity of interest is a remote radiation source or the properties of a propagation medium that perturbs the radiation), or sensors that provide data from which the quantity of interest must be inferred (indirect sensing) are not considered in this report and will be addressed in the next phase of the work. An example of indirect sensing is provided by a recent study of sensing of corrosion in metallic structures, and particularly at inaccessible locations such as in crevices or fastener holes, where damage cannot be measured directly but can be inferred from the moisture and other micro-climatic variables in the environment [Cole *et.al.*, 2008].

When dealing with direct measurements, one typically considers a sensor layout where sensors are placed over only a subset of possible locations, leaving “the rest of the space” without sensors. This is described by Guestrin *et al.* (2005) as: “not just interested in [sensor measurement] at sensed locations, but also at locations where no

sensors were placed”. However, measuring the quantity of interest in “the rest of the space” would still be a direct measurement.

## 1.2 Prior knowledge and environmental models

Finding an optimal sensor layout clearly requires knowledge (real or assumed) of the sensors and the environment. A sensor such as a thermocouple measures a quantity (temperature) at the location of the sensor: it is a local measurement. The density and locations of sensors required to map the temperature distribution in an environment depends on how this local measurement reflects the temperature in a region of the environment surrounding the thermocouple. This in turn depends on properties of the environment such as its physical structure, thermal properties and the nature and locations of any heat sources and sinks.

Three approaches to acquiring and applying knowledge of the environment to this problem are as follows.

1. A naïve approach to deciding the optimal placement of direct sensors in an environment is to assume the sensors have some fixed sensing radius (i.e. the sensor will indicate changes in the parameter of interest that occur within a fixed radius of the sensor position), and solve some variant of the art-gallery problem [González-Banos *et.al.*, 2001]. However, this is not realistic in practice as the sensing area is rarely a perfect circle [Guestrin *et.al.*, 2005].
2. Another approach involves learning a spatial model of the environment, such as a Gaussian process [Guestrin *et.al.*, 2005, Krause *et al.*, 2008] or a graphical model [Krause & Guestrin, 2005]. This approach may be most appropriate in situations where there is little *a priori* knowledge of the environment and its properties.
3. A third approach is to employ physical models of the sensor and environment. In this case the physical model encapsulates our prior knowledge of the structure or environment, or we can use a hybrid of approaches 2 and 3 by starting from an approximate physical model and refining its parameters by learning. This approach may be more appropriate for an engineered structure for which the properties are known.

There have been previous studies of optimal sensor placement for damage detection in structures based on physical models of the sensor-structure system [e.g. Staszewski *et al.*, 2000, Lee & Staszewski, 2007], but these have not utilised information-theoretic criteria for optimisation of the placements.

In some cases, however, important issues in engineered structures depend on unplanned and unpredictable features of the structure. For example, corrosion in well-designed and well-built structures such as aircraft may occur in places where defects in corrosion prevention seals and coatings have been introduced accidentally – perhaps near an ill-fitting seal, a joint that has been inadequately re-fitted or re-sealed after maintenance, fatigue in a fastener coating, etc. Such features will not generally be present in a physical model but may be detected by sensing. A hybrid approach that incorporates learning into the physical model may be more appropriate in such cases.

There is an important distinction between learnt statistical models, such as those employed in approach 2, and physical models (approach 3). Statistical models will generally be based on learning of sensor data, and will thus model the spatial dependence of sensor outputs over the environment of interest. They are data-driven

models, and will incorporate any effects of sensor noise and bias, which may subsequently influence the optimal sensor layouts derived from the models.

On the other hand, physical models of the environment will generally model the state of the environment, which is subsequently related to sensor data through a model of the sensor-environment interaction, which again may be either deterministic or statistical. The physical modelling approach would allow evaluation of the effects of sensor noise and bias on the results deduced from a sensor layout. Probabilistic sensor models are introduced in Section 2.2 below.

### **1.3 Sensor-environment scenarios**

Two different sensor-environment scenarios are discussed in this report. The first is a simple engineered structure, a thermal protection shield of a spacecraft, for which sensors are required to monitor effects of damage on its functional performance. Such a structure is expected to be amenable to the use of physical modelling (approach 3 above) to describe the sensor-environment properties. This scenario will be employed as an exemplar for description of the theoretical formulation introduced in the next sections, but we do not yet have experimental data. It is outlined in Section 2.1.

The second scenario is the measurement of soil moisture over an area of land on an agricultural property in northern Australia. Experimental data from an existing wireless sensor network has been obtained to enable testing and evaluation of the optimal sensor layout methods discussed and developed in this report. In this case there is relatively little knowledge of the relevant structure and properties of the environment (e.g. spatial distribution of soil types and structures, water transport patterns, etc.), so approach 2 has been adopted. A Bayesian Network model has been constructed, and parameter values have been learnt from a set of test or training data, as outlined in Section 4 below.

### **1.4 Cost and value**

As indicated above, it is important to find methods to place a limited number of sensor nodes in the area of interest such that the cost of the placement is minimised while the value of the obtained information is maximised. In this context, cost may include the installation cost, the energy cost of using the sensors, communication costs, the cost of processing and using the data, and any additional cost of operating the structure with the sensors deployed. Value is understood as the expected benefit of using the sensor configuration (or layout) for a given sensor model relative to the environment.

The cost and value of a sensor layout can be calculated with respect to the sensor-environment models using appropriate metrics. There is a multiplicity of metrics available in the current literature to compute the optimal sensor layout. These methods draw from experiment design [Ramakrishnan *et.al.*, 2005], decision theory [Krause & Guestrin, 2005a], and information theory [Guestrin *et.al.*, 2005, Olsson *et.al.*, 2004]. It is thus difficult to decide which metric is the most appropriate one, given a specific practical setup.

In this work we first introduce a criterion that maximises the value of using a sensor subset for direct measurements, given sensor and environment models. This value is defined as the difference between the optimal expected cost and the averaged optimal expected cost after utilising the sensors. Defining expected costs information-theoretically allows us to represent the sensor layout problem as an entropy

optimisation problem. This is then compared with other criteria, some of which are decision-theoretic while others are also information-theoretic. We compare the criteria presented using data from an existing wireless sensor network (the second scenario referred to in Section 1.3) by first computing the optimal configurations, then predicting the sensor measurements in the rest of the space and verifying the results using available ground truth.

This report is organised as follows. In Section 2, we describe a general approach to the optimal sensor layout problem. Section 3 briefly outlines several of the current methods in the literature. Section 4 presents the experimental setups of the sensor network and the environment model (i.e. graphical model of Bayesian Network) used for testing. Section 5 discusses the results of the experiment. Finally, Section 6 presents conclusions.

## 2. Formulation of the Optimal Design Problem

In this section, we formulate the sensor problem and present a general approach to optimal sensor placement. Arguably, taking observations is aimed at improving the outcome of a future action. Consider the decision problem where the cost incurred by a future decision is described by the function  $C : \mathcal{A} \times \mathcal{X} \rightarrow \mathfrak{R}$ , where  $\mathcal{A}$  is the set of possible decisions or actions,  $a$ , and  $\mathcal{X}$  is the set of possible world states,  $x$ , that are relevant to the decision problem. Thus,  $C(a, x)$  is the cost of carrying out action  $a$  when the system is in state  $x$ . There may be different costs for taking or failing to take a particular decision (to repair a defect, for example) when the environmental conditions required it. Note that this cost is different to that referred to in Section 1 above, which was the cost of deploying and operating a sensor: this is denoted by  $K$  (see Section 2.3). The value of the sensor deployment (Section 1) is the reduced cost of making an observation-assisted optimal action compared with that which would have been made without the assistance of the sensor.

If the true state  $x \in \mathcal{X}$  is known, the optimal action is easily found by:

$$a^* = \arg \min_{a \in \mathcal{A}} C(a, x) \quad (1)$$

However, for realistic systems, the state  $x$  is often not known precisely. For example, a sensor system may indicate the presence of damage in a structure but may not be capable of detailing the exact nature and extent of the damage: this uncertainty may be significant for a decision to repair or not repair the structure. Thus, it is necessary to consider the state to be the random variable  $X = [x_1, x_2, \dots, x_{|\mathcal{X}|}]$  with a given

probability distribution  $\mathcal{P}_x$  defined by  $\mathcal{P}_x = [P(x_1), P(x_2), \dots, P(x_{|\mathcal{X}|})]$ , where

$P(x) \equiv \Pr(X = x)$ . The distribution  $\mathcal{P}_x$  is called the *prior belief*, and defines a model of the environment in which the sensor is located. This notation assumes, for simplicity, that the state of the environment has discrete values, or can be characterised by discrete parameters.

Now, since the state is not known, the actual cost of a particular action cannot be determined with certainty and the expected cost of an action should be considered. The expected cost of an action is given by:

$$J(a, \mathcal{P}_x) = \sum_{x \in \mathcal{X}} P(x) C(a, x) \quad (2)$$



with the optimal expected cost given by:

$$J^*(\mathcal{P}_x) = \min_{a \in A} J(a, \mathcal{P}_x) \quad (3)$$

It is noted that this is a function of the given prior probability distribution  $\mathcal{P}_x$ . This generic characterisation does not include actual observations.

## 2.1 An illustrative scenario

For the sake of discussion and illustration of the formulation, a simple scenario will be adopted. This is not the scenario for which real sensor data will be analysed in later sections of this report (Sections 4 and 5): it is introduced here purely for discussion purposes.

It will be assumed that the monitored environment is the heat shield of a space vehicle. The property of interest is the thermal resistance shield, which may be reduced by damage, and which may be monitored by sensing the temperature on the inner surface when a known thermal source is applied to the external surface. Damage to the shield may be classified as negligible (0), non-critical (1) or critical (2). Negligible damage means that no temperature rise was detected when the thermal source was applied. Non-critical damage means some reduction in the thermal barrier property (i.e. a measurable temperature rise), but insufficient to put the vehicle in jeopardy during re-entry: repair will be required at the next scheduled maintenance. Critical damage means a sufficient reduction in the thermal protection to require immediate repair.

In this scenario we might have an initial prior belief that the shield has negligible damage, i.e.  $P(x=0)=1; P(x=1)=P(x=2)=0$ . However, if a potentially damaging event has been detected (e.g. an impact on the shield surface) some of these prior beliefs may be modified depending on the location and severity of the impact. A light impact, for example, might lead us to change the prior belief to, say,  $P(x=0)=0.2; P(x=1)=0.5; P(x=2)=0.3$ .

Possible actions can be defined as: do nothing ( $a=0$ ); schedule maintenance on return to base ( $a=1$ ); or immediate repair ( $a=2$ ). Other possibilities will be ignored in the interests of keeping the model simple. Associated with these actions there may be a cost matrix something like the following:

$$C = \begin{bmatrix} 0 & 0 & H \\ L & L & H \\ M & M & M \end{bmatrix}$$

where L, M, H indicate low, medium and high costs, and it is assumed that the cost of maintenance at base is low, that of repair in space is medium, and the cost of not immediately repairing a critical damage is high (perhaps the loss of the vehicle). Assigning numerical values for L, M and H will allow  $J$  and  $J^*$  to be calculated from equations (2) and (3) respectively. Note that  $C(a=0, x=1)$ , the cost of taking no action for a non-critical impact, has been assigned zero cost because that is the immediate outcome. However, longer-term costs may be significant if the damage develops further.

If the probabilities of the three damage states are  $P_0$ ,  $P_1$  and  $P_2$ , the expected costs of the three possible actions, from equation (2), are:

$$J(a=0) = P_2 H$$

$$J(a=1) = P_0 L + P_1 L + P_2 H$$

$$J(a=2) = P_0 M + P_1 M + P_2 M$$

The optimal expected cost (equation (3)) is the lowest of these, which will depend on the actual costs  $L$ ,  $M$  and  $H$ , and the damage probabilities. Clearly, if  $H$  is very much greater than  $M$ , then even a low probability of critical damage ( $P_2$ ) may lead to a high expected cost of not immediately repairing the damage.

Later it will be assumed that the shield can be considered to comprise a number  $M$  of small areas. These may be individual tiles, or small regions within a tile.

## 2.2 Inclusion of sensors and observations

Now consider a situation where a sensor can be deployed prior to taking the action to provide some information about the value of the state  $x$ . It is expected that this will lead to an improved final decision. Formally, the output of the sensor is denoted by the random variable  $Z$ , in the discrete domain  $\mathcal{Z}$ , i.e.  $Z = [z_1, z_2, \dots, z_{|\mathcal{Z}|}]$ , where the  $z_i$  are possible values of the sensor output.

The dependence of the sensor readings on the state  $x$  is modelled using a conditional probability distribution function.

$$\mathcal{P}_{Z|X} = \begin{bmatrix} P(z_1 | x_1) & \cdots & P(z_1 | x_{|\mathcal{X}|}) \\ \vdots & \ddots & \vdots \\ P(z_{|\mathcal{Z}|} | x_1) & \cdots & P(z_{|\mathcal{Z}|} | x_{|\mathcal{X}|}) \end{bmatrix} \quad (4)$$

where  $P(z | x) \equiv \Pr(Z = z | X = x)$ . In this notation, each value of  $x$  and  $z$  represents the set of values for the separate tiles. The distribution  $\mathcal{P}_{Z|X}$  is called the *sensor model* and completely specifies the characteristics of the sensor.

After a specific measurement  $z$  has been made, a new conditional distribution  $\mathcal{P}_{X|z}$ , defined as  $[P(x_1 | z), P(x_2 | z), \dots, P(x_{|\mathcal{X}|} | z)]$ , where  $P(x | z) \equiv \Pr(X = x | Z = z)$ , for the value of  $X$  can be generated using Bayes rule:

$$P(x | z) = \frac{P(z | x)P(x)}{\sum_{x \in \mathcal{X}} P(z | x)P(x)} \quad (5)$$

The distribution  $\mathcal{P}_{X|z}$  will be referred to as the *posterior belief* and is defined on the same domain as  $\mathcal{P}_X$ . The optimal action, for this belief, has an expected cost of

$$J^*(\mathcal{P}_{X|z}) = \min_{a \in \mathcal{A}} \sum_{x \in \mathcal{X}} P(x | z) C(a, x) \quad (6)$$

This is a *posterior* measure, since it requires the value of the observation  $z$ . An *a priori* measure can be constructed by considering the expectation over all observations  $z \in \mathcal{Z}$  of  $J^*(\mathcal{P}_{X|z})$  and a distribution over the observations,  $P(z)$ :

$$\begin{aligned}
G^*(\mathcal{P}_X, \mathcal{P}_{Z|X}) &= \sum_{z \in Z} P(z) J^*(\mathcal{P}_{X|z}) \\
&= \sum_{z \in Z} P(z) \min_{a \in A} \left\{ \sum_{x \in X} P(x|z) C(a, x) \right\}.
\end{aligned} \tag{7}$$

The function  $G^*$  is completely specified by the prior belief,  $\mathcal{P}_X$ , and the sensor model,  $\mathcal{P}_{Z|X}$ .

Now, if the function  $J^*$  is concave (for example, if  $J^*(Y) = \sum_{y \in Y} y \log \frac{1}{y}$ , for some array  $Y$ ), then Jensen's inequality gives us:

$$G^*(\mathcal{P}_X, \mathcal{P}_{Z|X}) = \sum_{z \in Z} P(z) J^*(\mathcal{P}_{X|z}) \leq J^*\left(\sum_{z \in Z} P(z) \mathcal{P}_{X|z}\right) = J^*(\mathcal{P}_X) \tag{8}$$

This means that  $G^*(\mathcal{P}_X, \mathcal{P}_{Z|X})$  is never greater than  $J^*(\mathcal{P}_X)$ , or put another way, on average utilising a sensor is never detrimental to the final decision problem, irrespective of the sensor model. We will prove later that the function  $J^*$  is concave. The difference between  $G^*$  and  $J^*$  can be used to determine the value or expected benefit of using the sensor modelled by  $\mathcal{P}_{Z|X}$  for a given prior belief  $\mathcal{P}_X$ :

$$V(\mathcal{P}_X, \mathcal{P}_{Z|X}) = J^*(\mathcal{P}_X) - G^*(\mathcal{P}_X, \mathcal{P}_{Z|X}) \geq 0. \tag{9}$$

Here, the value of using a sensor with a given model  $\mathcal{P}_{Z|X}$  is explicitly dependent on the information that is already available, defined by  $\mathcal{P}_X$ .

### 2.3 Optimal sensor layout

The question that this paper attempts to explore is, if there are multiple sensors (and layouts) available, each associated with a different sensor model, which is the best to use?

We consider a general case of problems where the state of interest is defined as a set of variables  $\mathbf{X} = \{X^1, X^2, \dots, X^m\}$ , where the individual arrays of state values  $X^i$  may refer to sub-regions of the environment or structure or perhaps to different properties of the environment.

If there are  $n$  sensors, their outputs will be described by the random variables  $\mathbf{Z} = \{Z^1, Z^2, \dots, Z^n\}$ , where each has an associated sensor model  $\mathcal{P}_{Z^i|\mathbf{X}}$ . [If sensor  $i$  measures only the state associated with state variable  $j$ , this sensor model will reduce to  $\mathcal{P}_{Z^i|X^j}$ . This assumption will be introduced later (Section 2.6).]

The value of using a particular sensor  $i \in \{1, \dots, n\}$  is given by its expected value, defined in (9). Then, if the  $i^{\text{th}}$  sensor has a cost  $K(i)$ , the optimal sensor can be found by simply maximising the net benefit, for the given prior belief  $\mathcal{P}_X$ :

$$i^* = \arg \max_i V(\mathcal{P}_X, \mathcal{P}_{Z^i|\mathbf{X}}) - K(i) \tag{10}$$

This formulation explicitly captures the prior information  $\mathcal{P}_{\mathbf{X}}$ , the usage cost  $K(i)$  and the characteristics of the sensor model  $\mathcal{P}_{\mathbf{Z}|\mathbf{X}}$ .

A reason why this process is not often used for sensor layout design tasks (such as where to place temperature sensors in a building) is the difficulty it introduces in the formulation of the final decision problem (what is the temperature information going to be used for?). In other words, the function  $C(a, x)$  is often not known in advance. This problem is typically overcome by ignoring the final decision task and simply reformulating the problem of optimal sensor layout as an inference problem that selects the sensor layout which maximises the information collected by the sensors about the variable array  $\mathbf{X}$ . This will be described in the next section.

## 2.4 Information-theoretic measures

Information theory provides the tools required to quantify what is meant by “information collected by the sensors”. Entropy provides a measure of the uncertainty associated with a belief  $\mathcal{P}_{\mathbf{X}}$ :

$$H(\mathbf{X}) = \sum_{x \in \mathcal{X}} P(x) \log \frac{1}{P(x)} \quad (11)$$

This was first derived by Shannon (1948) from 3 basic axioms that a metric of uncertainty should satisfy. The base of the logarithm determines the unit of the measure, with base 2 corresponding to “bits”.

Thus, to formulate the sensor selection problem as an information maximisation problem, the optimal expected cost  $J^*$  of a distribution is replaced by the entropy  $H$ , that is  $J^*(\cdot) = H(\cdot)$  and  $J^*(\cdot | \cdot) = H(\cdot | \cdot)$ . Further, the entropy is a concave function, so the inequality in equation (8) is satisfied. Intuitively, the information maximisation criterion suggests the use of the sensor that, on average, produces the least uncertain posterior belief. It is noted that these two different criteria (generic and information maximising) may select different optimal sensors.

Substituting  $J^*(\cdot)$  with  $H(\cdot)$  in (7) transforms the function  $G^*(\mathcal{P}_{\mathbf{X}}, \mathcal{P}_{\mathbf{Z}|\mathbf{X}})$  into the conditional entropy of  $\mathbf{X}$  given  $\mathbf{Z}$ ,

$$G^*(\mathcal{P}_{\mathbf{X}}, \mathcal{P}_{\mathbf{Z}|\mathbf{X}}) = \sum_{z \in \mathcal{Z}} P(z) H(\mathbf{X} | z) \quad (12)$$

$$= \sum_{z \in \mathcal{Z}} P(z) \sum_{x \in \mathcal{X}} P(x | z) \log \frac{1}{P(x | z)} \quad (13)$$

$$= H(\mathbf{X} | \mathbf{Z}) \quad (14)$$

where  $\mathcal{X}, \mathcal{Z}$  are now understood as the domains of the variable sets  $\mathbf{X}$  and  $\mathbf{Z}$ , and the value of a configuration becomes the mutual information between  $\mathbf{X}$  and  $\mathbf{Z}$ ,

$$V(\mathcal{P}_{\mathbf{X}}, \mathcal{P}_{\mathbf{Z}|\mathbf{X}}) = J^*(\mathcal{P}_{\mathbf{X}}) - G^*(\mathcal{P}_{\mathbf{X}}, \mathcal{P}_{\mathbf{Z}|\mathbf{X}}) \quad (15)$$

$$= I(\mathbf{X}; \mathbf{Z}) \quad (16)$$

If each sensor has a different usage cost, the optimal selection problem becomes ill-posed since there is no direct method of trading off information with usage cost without explicitly considering what the information will be used for (an information theoretic approach was used to avoid this). To overcome this problem it will be

assumed that the usage costs are constant and the optimal design problem can be formulated as the maximisation, over all sensors  $i$ , of the mutual information between  $\mathbf{X}$  and  $Z^i$

$$i^* = \arg \max_i I(\mathbf{X}; Z^i) \quad (17)$$

Alternatively, we can convert the information maximisation problem into an entropy minimisation problem when selecting the best sensor:

$$i^* = \arg \max_i I(\mathbf{X}; Z^i) \quad (18)$$

$$= \arg \max_i [H(\mathbf{X}) - H(\mathbf{X} | Z^i)] = \arg \min_i H(\mathbf{X} | Z^i) \quad (19)$$

Note that the equivalence between (18) and (19) does not scale to the selection of multiple sensors (see Section 3.1).

## 2.5 Graphical representations of information theoretic quantities

A useful graphical representation of the relationships between different information-theoretic functions of the variables is the  $I$  diagram [Yeung, 1991], which is similar to the Venn diagram in set theory<sup>1</sup>. Figure 1 shows the relationships between the entropies of  $X$  and  $Z$ , and their mutual information  $I$ . Each area in this figure represents an amount of entropy or uncertainty. For example, the blue area labelled  $H(X | Z)$  represents the average uncertainty remaining about  $X$  after the sensor data  $Z$  has been obtained. The overlap in the middle (green) represents the mutual information between the variables. The circular region labelled  $H(X)$  (i.e. blue + green) is the uncertainty in the state variable  $X$ .

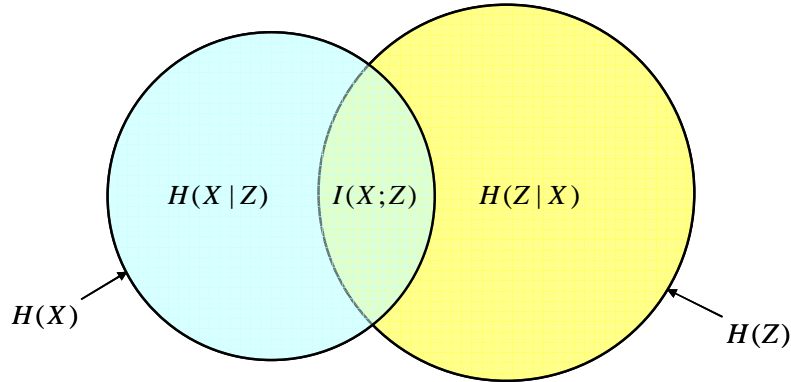


Figure 1:  $I$  diagram for the random variables  $X$  and  $Z$ .

For a sensor design problem it is the overlap region  $I(X;Z)$  that should be maximised, or conversely the residual uncertainty  $H(X | Z)$  minimised. The maximum value of the mutual information is the smaller of  $H(X)$  and  $H(Z)$ .

If two sensors are now considered, and there are three random variables  $X, Z^1$  and  $Z^2$ , the situation may be as shown in Figure 2. In this case  $H(X | Z^1) < H(X | Z^2)$  and  $I(X;Z^1) > I(X;Z^2)$  as depicted.

<sup>1</sup> However, unlike a Venn diagram some areas may represent negative quantities [Yeung, 1991, MacKay, 2003].

This representation allows the optimisation criterion, defined in equations (17) and (19), to be illustrated graphically. From the  $I$  diagram (Figure 2) it is clear that these two formulations of the criterion are the same, as derived algebraically.

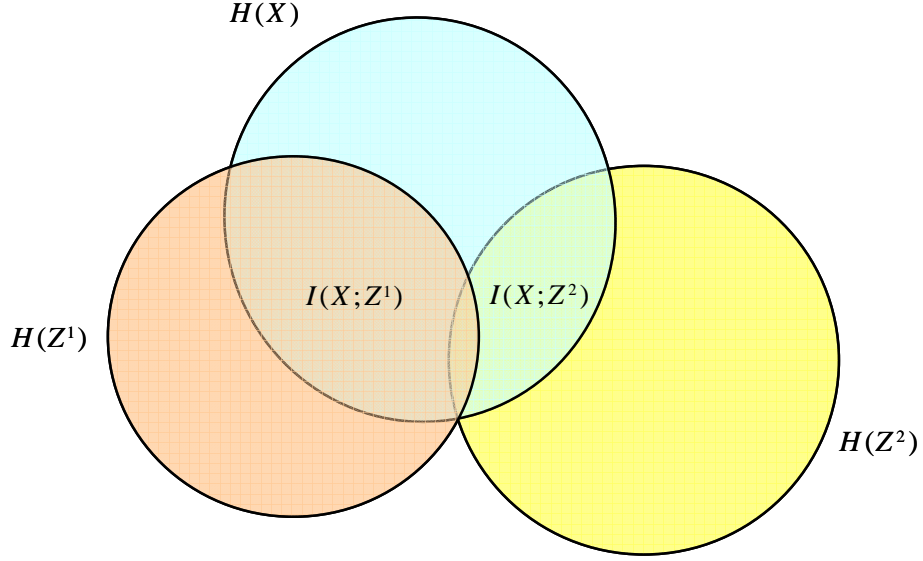


Figure 2:  $I$  diagram for the random variables  $X$ ,  $Z^1$  and  $Z^2$ .

## 2.6 Optimal selection of a subset

At this stage, we apply this approach to direct measurements. In the simplest case of selecting one optimal sensor  $\mathbf{Z}$ , which is a deterministic function of the state  $\mathbf{X}$ , the conditional entropy of  $\mathbf{Z}$  given  $\mathbf{X}$  must be zero. This occurs because once the state  $\mathbf{X}$  is known there is no uncertainty in  $\mathbf{Z}$  (Figure 3).

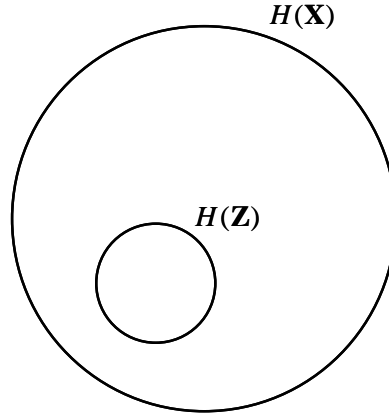


Figure 3:  $I$  diagram for a system where the sensor observations are a deterministic function of the variable  $\mathbf{X}$ .

It is clear then that the mutual information,  $I(\mathbf{X}; \mathbf{Z})$ , between the observation  $\mathbf{Z}$  and the state  $\mathbf{X}$  is equal to the entropy of the observation  $H(\mathbf{Z})$ . Using this in the sensor selection criterion (17) yields the optimal sensor as follows:

$$i^* = \arg \max_i I(\mathbf{X}; Z^i) \quad (20)$$

$$= \arg \max_i H(Z^i) \quad (21)$$

We now consider a slightly more specialised case of problems where the number of sensors is the same as the number of state variables (i.e.  $n = m$ ). Thus, the state of interest  $\mathbf{X}$  is defined as a set of variables  $\mathbf{X} = \{X^1, X^2, \dots, X^m\}$  and it is assumed that a set of sensors  $\mathbf{Z} = \{Z^1, Z^2, \dots, Z^m\}$  exists, where each can measure the value of an associated variable  $X^i$ , which is the case of direct measurement as discussed earlier. This may be a thermal protection shield that is divided into  $m$  incremental areas (perhaps individual tiles), each of which contains a temperature sensor, or, in our second scenario (Section 1.3) a sub-region monitored by a single moisture sensor.

The design task becomes to select a subset of sensors  $\nu \subseteq \{1, 2, \dots, m\}$  to deploy. To avoid the case of selecting all variables  $\nu = \{1, 2, \dots, m\}$ , a constraint is generally imposed on  $\nu$ . For simplicity it is assumed this constraint imposes a maximum limit  $r$  on the number of elements in  $\nu$ . To be able to refer to the elements of  $\nu$  explicitly, this set will be denoted by  $\nu = \{i_1, i_2, \dots, i_r\}$ .

Now consider an abstract compound random variable  $\mathbf{Z}^\nu$ , representing the combined output of all selected variables and defined as:

$$\mathbf{Z}^\nu = \{X^i : \forall i \in \nu\} = \{X^{i_1}, X^{i_2}, \dots, X^{i_r}\}.$$

This notation specialises the problem to one of direct measurement with noiseless, unbiased sensors, i.e. one in which the sensor data gives a direct measure of the state variable. Whether or not the sensors are noiseless and unbiased (which will generally not be the case) this assumption will still be a reasonable one for statistical data-driven sensor-environment models (case 2 in Section 1.2). This is because such models learn from sensor data rather than from state information, and thus are models that represent the spatial distribution of sensor outputs rather than of environmental state variables. The experimental case discussed and analysed in Section 4 utilises such a model, which is the reason this notation has been introduced here.

Thus the optimal sensor selection task now becomes:

$$\nu^* = \arg \max_{|\nu| \leq r} I(\mathbf{X}; \mathbf{Z}^\nu) \quad (22)$$

$$= \arg \max_{|\nu| \leq r} I(X^1, X^2, \dots, X^m; \mathbf{Z}^\nu) \quad (23)$$

$$= \arg \max_{|\nu| \leq r} H(\mathbf{Z}^\nu) \quad (24)$$

### 3. Current Methods

The current literature discusses several other optimal sensor placement methods, three of which will now be outlined.

#### 3.1 Reward and entropy (RE)

Krause and Guestrin (2005a) introduced a *local reward function*  $R$ , which is defined on the marginal probability distribution of the variables in  $\mathbf{X}$ . The local reward is set for each variable  $X^i$  as the conditional entropy given the observation variable  $\mathbf{Z}^\nu$ , i.e.:

$$R_i(P(X^i | \mathbf{Z}^\nu)) = -H(X^i | \mathbf{Z}^\nu) \quad (25)$$

The objective of the optimisation then becomes the minimisation of the sum of conditional entropies:

$$\nu_R^* = \arg \min_{|\nu| \leq r} \sum_{i=1}^m H(X^i | \mathbf{Z}^\nu) \quad (26)$$

This will be referred to later as the RE criterion. It is noted that in general this is not the same criterion as the one developed in the previous section, since it does not take into account dependencies between the variables  $X^i$ . This can be demonstrated by comparing with equation (23):

$$\begin{aligned} \arg \max_{|\nu| \leq r} I(X^1, \dots, X^m; \mathbf{Z}^\nu) &= \arg \max_{|\nu| \leq r} \left[ \underbrace{H(X^1, \dots, X^m)}_{\text{Constant}} - H(X^1, \dots, X^m | \mathbf{Z}^\nu) \right] \\ &= \arg \min_{|\nu| \leq r} H(X^1, \dots, X^m | \mathbf{Z}^\nu) \\ &= \arg \min_{|\nu| \leq r} \left[ \sum_{i=1}^m H(X^i | \mathbf{Z}^\nu) - \underbrace{I(X^1; X^{2:m} | \mathbf{Z}^\nu) - \dots - I(X^{m-1}; X^m | \mathbf{Z}^\nu)}_{\text{Terms not accounted for in (26)}} \right] \end{aligned}$$

### 3.2 Mutual information (MI)

Guestrin *et al.* (2005) proposed a metric that gives an optimum subset of sensor locations that minimises the uncertainty about the estimates in the “rest of the space”. The problem is formulated by searching for  $\mathbf{Z}^\nu$  that reduces the entropy over the rest of the space  $\mathbf{X} \setminus \mathbf{Z}^\nu = \mathbf{X}^{\bar{\nu}} = \{X^i : i \notin \nu\}$ . Formally, the optimal subset is:

$$\begin{aligned} \nu_{\text{MI}}^* &= \arg \max_{|\nu| \leq r} H(\mathbf{X}^{\bar{\nu}}) - H(\mathbf{X}^{\bar{\nu}} | \mathbf{Z}^\nu) \\ &= \arg \max_{|\nu| \leq r} I(\mathbf{X}^{\bar{\nu}}; \mathbf{Z}^\nu) \end{aligned} \quad (27)$$

Thus this measure is equivalent to finding the maximum mutual information between  $\mathbf{X}^{\bar{\nu}}$  and  $\mathbf{Z}^\nu$ , and the criterion will be referred to below as MI.

This only takes into account the mutual information between the observed and unobserved variables and not the remaining uncertainty of the unobserved variables. As for the RE criterion, this can be demonstrated by comparing with equation (23):

$$\begin{aligned} \arg \max_{|\nu| \leq r} I(X^1, \dots, X^m; \mathbf{Z}^\nu) &= \arg \max_{|\nu| \leq r} \left[ \underbrace{H(X^1, \dots, X^m)}_{\text{Constant}} - H(X^1, \dots, X^m | \mathbf{Z}^\nu) \right] \\ &= \arg \max_{|\nu| \leq r} [-H(X^1, \dots, X^m | \mathbf{Z}^\nu)] \\ &= \arg \max_{|\nu| \leq r} \left[ -H(\mathbf{X}^{\bar{\nu}} | \mathbf{Z}^\nu) - \underbrace{H(\mathbf{X}^\nu | \mathbf{Z}^\nu)}_{=0} + \underbrace{I(\mathbf{X}^{\bar{\nu}}; \mathbf{X}^\nu | \mathbf{Z}^\nu)}_{=0} \right] \\ &= \arg \max_{|\nu| \leq r} [H(\mathbf{X}^{\bar{\nu}}) - H(\mathbf{X}^{\bar{\nu}} | \mathbf{Z}^\nu) - H(\mathbf{X}^{\bar{\nu}})] \\ &= \arg \max_{|\nu| \leq r} \left[ I(\mathbf{X}^{\bar{\nu}}; \mathbf{Z}^\nu) - \underbrace{H(\mathbf{X}^{\bar{\nu}})}_{\text{Term not accounted for in (27)}} \right] \end{aligned}$$



This work has been extended by Krause *et al.* (2008). Trendafilova *et al.* (2001) also utilised a mutual information criterion for sensor layouts for damage detection. In this case the aim was to minimise the average mutual information between sensor data (using data from accelerometers distributed on a vibrating plate), to ensure minimum redundancy between the measurements of the set of sensors.

### 3.3 Information coverage (IC)

Olsson *et al.* (2004) suggested a method that uses a combination of mutual information and an information metric [Crutchfield, 1990] to describe an *information coverage* criterion:

$$\nu_{IC}^* = \arg \max_{|V| \leq r} \sum_{i \in V} \sum_{\substack{j \in V \\ i \neq j}} [w_1 I(Z^i; Z^j) + w_2 (H(Z^i | Z^j) + H(Z^j | Z^i))] \quad (28)$$

where the mutual information  $I(Z^i; Z^j)$  is used as a measure for redundancy between measurements  $Z^i$  and  $Z^j$ , and the information metric (i.e. the information distance between two sensors),  $H(Z^i | Z^j) + H(Z^j | Z^i)$ , is used as a measure for novelty,  $w_1$  and  $w_2$  are the weights used to emphasise redundancy and novelty. Greater redundancy in this case improves the robustness of the sensors against noise in the environment. Novelty is the measure that captures as much different information as possible from the environment. It is noted that neither the model of the environment,  $\mathcal{P}_x$ , nor the sensor models,  $\mathcal{P}_{z|x}$ , are used in this approach, which will be referred to as IC below.

If we set  $w_1 = w_2 = 1$ , then equation (28) reduces to:

$$\nu_{IC}^* = \arg \max_{|V| \leq r} \sum_{i \in V} \sum_{j \in V} [H(Z^i | Z^j)] \quad (29)$$

Thus, it is important to choose the ratio between  $w_1$  and  $w_2$ , to capture redundancy or novelty in the system. Following Olsson *et al.* (2004), in the analysis presented in Section 5, we used  $w_1 = 1$  and  $w_2 = 4$  to put more emphasis on the novelty, but these values can be varied arbitrarily.

## 4. Experimental Setup

This section describes an experiment in which real network data from soil moisture measurements is used to derive optimal sensor placements using the four criteria outlined above: the criterion introduced here that reduces to equation (24), and the RE, MI and IC criteria outlined above. Data from sensors on an approximately rectangular 4x4 grid is used to find the optimal locations on the grid if only 2, 3 or 4 sensors were used. The resulting sensor layouts are tested using another set of data, obtained from a different time.

### 4.1 Data

The data set used for this paper is obtained from a current wireless sensor network in Belmont, Australia, some 670 km north of Brisbane. The network was set up as a test bed for environmental and animal behaviour monitoring at Belmont Research Station. The fixed environmental nodes used for this paper are solar powered and has onboard sensors for monitoring soil moisture, battery voltage, and solar voltage. Figure 4

shows the configuration of the fixed nodes, the numbers in brackets are the replacement nodes.

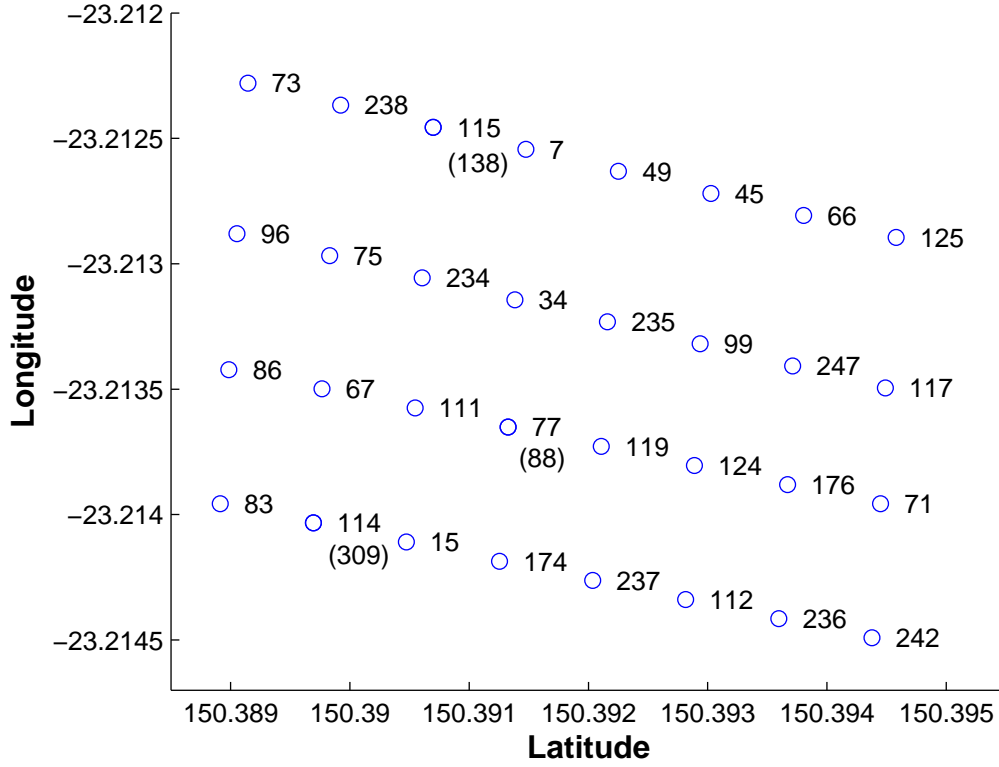
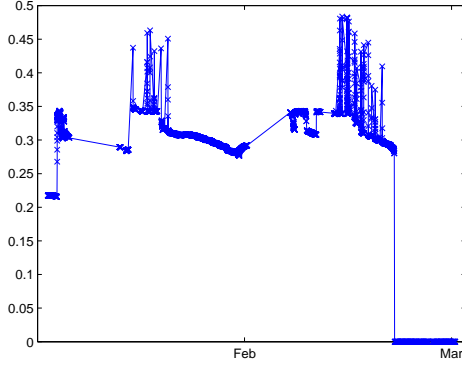
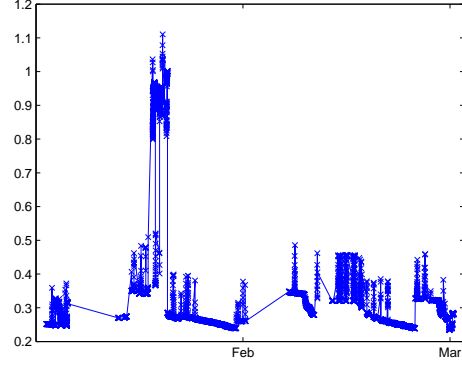


Figure 4: The configuration of the sensor network used described in this section. The numbers shown are the sensor ID numbers, and the numbers in brackets are replacement sensors introduced during the period the data was taken.

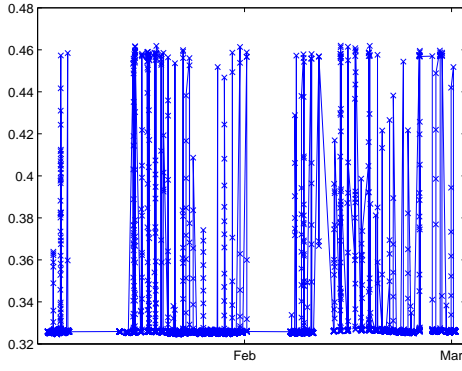
We used approximately two months of the soil moisture data from January and February 2008. Each node takes a sensor reading at roughly one minute intervals, independent of its neighbours. The data are preprocessed such that all the readings occur at the same time on 0 seconds of a minute. Further, due to various environmental and onboard issues, some nodes may not record any data for a period of time. The irregularity in the individual sensor's data time stamps combined with drop-out in data recording of individual sensors means not all sensor nodes in the network will record a reading at a given time stamp  $t$ . In other words, the data set includes missing values. The data recorded by the first four nodes in the network is shown in Figure 5.



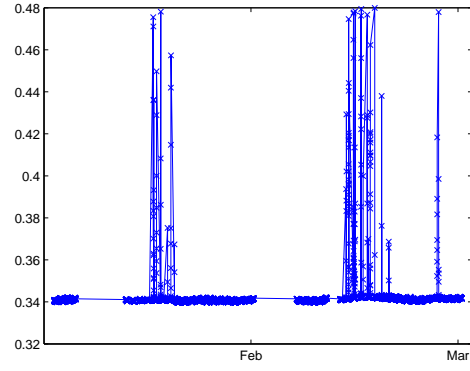
(a) Sensor ID 73



(b) Sensor ID 96



(c) Sensor ID 86



(d) Sensor ID 83

Figure 5: The soil moisture data collected by the first four sensors in the network during the months of January and February 2008. The ‘x’ marks are valid readings, and the y-axis shows raw data collected by the sensor. The gaps between valid data points represent periods when no data was communicated by the sensor.

## 4.2 Bayesian networks

As indicated in Section 1, a graphical model, specifically a Bayesian Network [Wang *et.al.*, 2008], was employed in this case to model the environment,  $\mathcal{P}_{\mathbf{x}}$ . Figure 6 shows the structure of the Bayesian Network (BN) used. Only the nodes west of  $150.392^\circ$  latitude (i.e. one half, see Figure 4), of the sensor network were used in constructing the spatial model, i.e. data from the western half of the network was used as learning data to construct the model. Data from the same set of sensors for a later time period was subsequently used for the comparison of optimal sensor layout criteria, which will be discussed in the next section.

The network was constructed using the assumption that neighbouring nodes in the sensor network are interdependent. Specifically, each node  $X^i$  in the BN is a parent to two neighbouring nodes, and a child to two neighbouring nodes. The joint distribution of the spatial model as described by the BN is:

$$P(\mathbf{X}) = P(X^1) \cdot \prod_{k1} P(X^{k1} | X^{k1-1}) \cdot \prod_{k2} P(X^{k2} | X^{k2-4}) \cdot \prod_{k3} P(X^{k3} | X^{k3-4}, X^{k3-1}), \quad (30)$$

where  $k1 = \{2,3,4\}$ ,  $k2 = \{5,9,13\}$ , and  $k3 = \{6,7,8,10,11,12,14,15,16\}$  (see Figure 6).

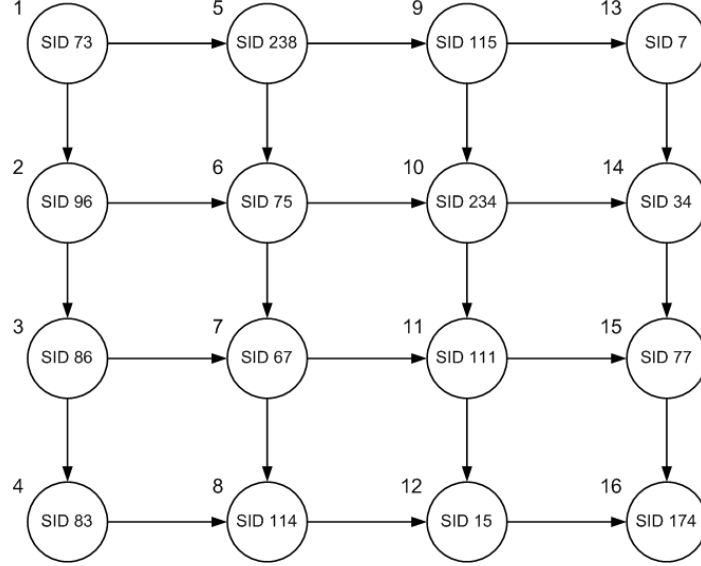


Figure 6: The Bayesian Network (BN) model used for learning the sensor network. The numbers on the top left side of each node denotes the BN node number.

### 4.3 Learning and inference

In a Bayesian Network the aim of the learning process is to estimate the parameters as well as to find the structure of the network. The objective in the learning is to find a network that “best describes” the probability distribution over the training data [Pearl, 1988]. In this work, however, the structure of the network was assumed to be known, and only the parameters needed to be learnt. The Maximum Likelihood [MacKay, 2003, Myung, 2003] algorithm could not be used in this case since the data contains hidden values, that is, each sensor node has not recorded sensor readings at every time stamp  $t$ . Therefore, the Expectation Maximisation (EM) algorithm [Dempster *et.al.*, 1977, Cowell *et.al.*, 1999] was used. The EM algorithm provides a general approach to maximum-likelihood parameter estimation when training data is incomplete.

This learnt network can then be used to carry out inference tests on new data. That is, given the observed values of some of the nodes in the network, compute the probability distribution of the data for other nodes. Inference allows us to perform *prediction* on the data, that is, the posterior probability distribution of the child node can be computed given the values of the parents. The prediction results are then compared with the ‘ground truth’ measured test data to compare the performances of the various sensor layout metrics.

As described in Section 4.1 the data set was divided into two, corresponding to nodes in the western and eastern halves of the network. The western half was used for training and, for data from a subsequent time period, for testing. We further processed the data to give a discretisation of 3 values, {low, median, high}, to be used in the discrete nodes. No other pre-processing was carried out.

## 5. Results and Discussion

This section presents the results of our experiments aimed at finding optimal layouts for two, three and four sensors respectively, using the methods and criteria described earlier. For each subsection, the optimal layout is first presented, followed by the inference results for the other nodes in the network. The learnt Bayesian Network was used to find the optimal layout of the sensors by finding each probability distribution,  $p(\cdot)$ , through marginalisation of the joint distribution. A greedy search process was employed, thus giving the best theoretically possible results for the various criteria.

### 5.1 Two Sensors

Figure 7 shows the results of optimal layout for two sensors using the four different criteria. Each image shows the value of the respective criterion for every combination of locations for two sensors. Figure 5(a) shows the results of using the  $H(\mathbf{Z})$  criterion of Section 2.6. It can be seen that the cells (1, 13) and (13, 1) have the highest entropy, that is, combining sensor locations at node 1 and node 13 is the best choice according to this criterion. Figure 4 shows that these two sensors are located at the top two corners of the Bayesian Network, which seems to confirm the proposition of Guestrin *et.al.* (2005) that the entropy-based method “pushes” the sensors to the edges of the network. Similarly, Figure 7(d) shows that using information coverage gives the same optimal layout even though not all cells agree between Figures 7(a) and (d).

Figure 7(b) shows the results of using the RE criterion (Section 3.1). In this case, the optimum layout is obtained by finding the minimum of all the values, which is given by the combination of nodes 1 and 10. This combination has one node at the top left corner and the other near the middle. Figure 7(c) shows the results of using the MI criterion (Section 3.2). The optimum layout here is obtained by finding the maximum of all the values, giving the combination of nodes 2 and 6, which are the first and second cells of the second row, thus one on the edge and the other near the middle of the array. Table 4, first row, summarises the results of the optimum placement for 2 sensors.

These sensor placements were evaluated by performing inferences on the rest of the nodes, i.e.  $\mathbf{X} \setminus \mathbf{Z}^v$ , to compare their prediction results with the actual measurements (ground truth). The results are represented in the form of a confusion matrix [Fawcett, 2003] because it gives an intuitive representation of the prediction performance. Since the data values are discrete, prediction in this case is similar to classification, because the considered prediction values can be interpreted as class labels. However, strictly speaking, a prediction task has been performed rather than a classification task. The prediction results were evaluated by counting the number of correct predictions for every discrete value.

Table 1 presents results for the prediction results of  $\mathbf{X} \setminus \mathbf{Z}^v$  given  $\mathbf{Z}^v = \{1, 13\}$ , the optimal layout obtained by using both the  $H(\mathbf{Z})$  and the IC criteria. The number of ground truth data samples for each value is shown in column 2. It can be seen that there is a near-perfect prediction result for all three values, which suggests that the sensors placed at the corners of the network do provide enough information to infer the possible sensor measurement at other locations.

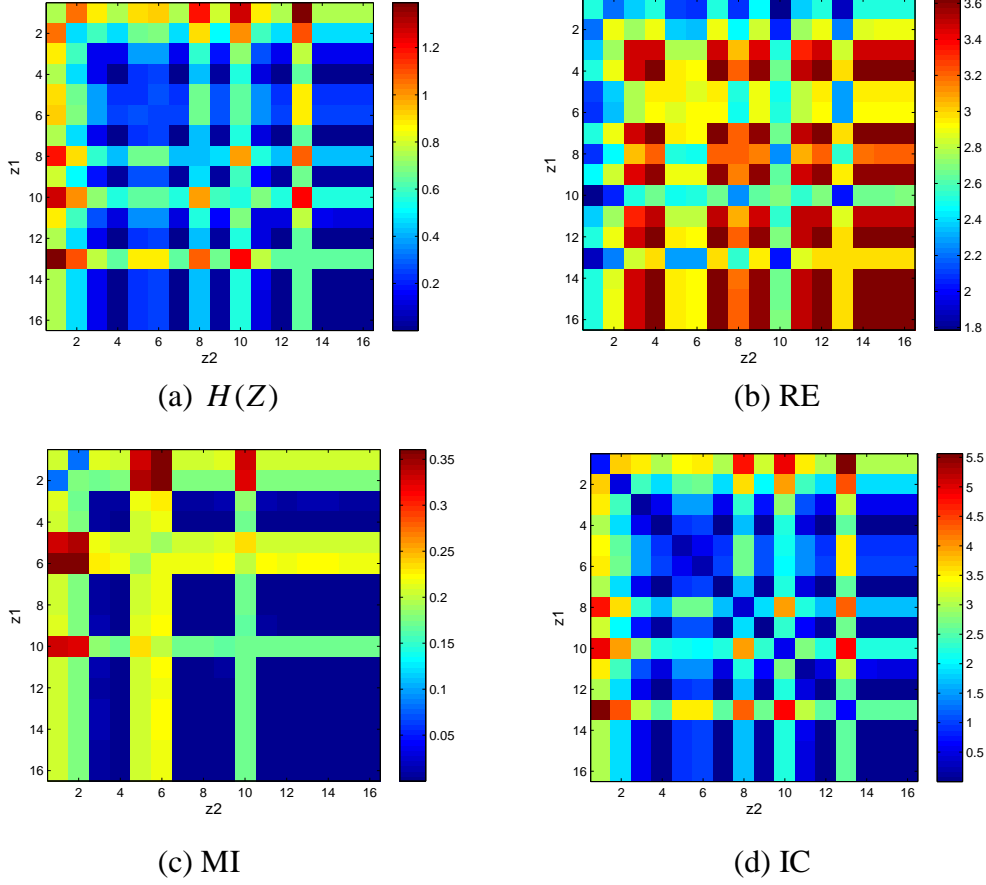


Figure 7: The results of optimal layout for two sensors for the four different criteria. For (a), (c) and (d), the optimal layout is the combination of sensor locations that gives the maximum value; for (b) it is the one with the minimum value.

The number of valid measurements with which the predicted values are compared differ for the different nodes because of sensor drop-outs, as indicated above. For this reason the sum of the  $n$  values in the following tables differ significantly for different sensor combinations.

Table 1: Inference results for sensor combination at node 1 and node 13.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	5542	<b>100.0</b>	0	0
2	12536	0	<b>99.52</b>	0.48
3	11220	0	1.78	<b>98.22</b>

Table 2 shows the prediction results where the optimal layout as suggested by the RE criterion. Comparing with the results in Table 1, it can be seen that there is a slight reduction in prediction performance for all three values. A two-tailed hypothesis test was used to compare these results statistically, taking the different sizes of the ground truth data into account. The null hypothesis  $H_0$  is set to be the hypothesis that the results in Table 2 were from the same distribution as those results in Table 1 and the

$\alpha$ -value was set at 0.05. The resulting  $P$ -values<sup>1</sup> for comparison of the three values were found to be all less than  $1 \times 10^{-20}$ , much smaller than the  $\alpha$  value, which means the observed differences are significant, and thus the null hypothesis can be rejected. Therefore, the sensor combination  $\mathbf{Z}^\nu = \{1, 10\}$  deduced from the RE criterion has slightly worse performance than the sensor combination  $\mathbf{Z}^\nu = \{1, 13\}$ .

Table 2: Inference results for sensor combination at node 1 and node 10.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	5851	<b>91.06</b>	8.94	0
2	10832	0	<b>99.33</b>	0.67
3	11379	0	2.02	<b>97.98</b>

Table 3 shows the prediction results given  $\mathbf{Z}^\nu = \{2, 6\}$ , which is the optimal layout results from the MI criterion. Comparing with the results shown in Table 1, it can be seen that there is a large drop in prediction performance for value 1, and a small decrease in value 3. Using two-tailed hypothesis tests and setting  $H_0$  to be similar as before, and  $\alpha = 0.05$  again, all  $P$ -values were again found to be near 0, and thus all differences were significant. However, although setting the two sensors at  $\mathbf{Z}^\nu = \{2, 6\}$  gives slightly better performance for value 2, it has much worse performance for values 1 and 3 than those results from setting sensors at  $\mathbf{Z}^\nu = \{1, 13\}$ .

Table 3: Inference results for sensor combination at node 2 and node 6.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	5928	<b>78.93</b>	21.07	0
2	9449	0	<b>99.57</b>	0.43
3	9309	0	6.03	<b>93.97</b>

The observation that the criterion deduced in this work (Section 2.6) produced very similar results to the information coverage criterion of Olsson *et al.* (2004) is interesting and requires further investigation to find under what conditions this similarity occurs. It would seem unlikely to be generally the case in view of the arbitrary choice of redundancy and novelty values introduced in the IC criterion, and also because it does not depend on prior knowledge or the sensor model.

Table 4: Results of optimal layouts for multiple sensors

No. of sensors	$H(\mathbf{Z})$ criterion	RE criterion	MI criterion	IC criterion
2	1, 13	1, 10	2, 6	1, 13
3	1, 10, 13	1, 10, 13	1, 3, 10	1, 10, 13
4	1, 8, 10, 13	1, 8, 10, 13	2, 5, 10, 13	1, 8, 10, 13

<sup>1</sup> The  $P$ -value, or significance value, is the probability of observing the test statistic if the null hypothesis is true.

## 5.2 Three and Four Sensors

Similar heuristic searches were performed to find the optimal layouts with three sensors using the four different criteria, and the results are summarised in Table 4. In this case, there are only two different configurations given by the four criteria:  $\mathbf{Z}^\nu = \{1,3,10\}$  for the MI criterion and  $\mathbf{Z}^\nu = \{1,10,13\}$  for the other three. Moreover, all resulting layouts share the nodes 1 and 10, and the latter is near the middle of the network. This does not agree with the conclusion by Guestrin *et al.* (2005) that using the entropy criterion will result in sensors being placed far apart along the boundary of the space. Further, the results from the optimal layouts with four sensors, shown also in Table 4, indicates that this is not an accidental case. It is thought to be related to the use of continuous rather than discrete variables by Guestrin *et al.* (2005), leading to a different (incorrect) definition of entropy.

The prediction results from these two sensor layouts deduced for three sensors are shown in Tables 5 and 6. Comparing the results, it can be seen that the sensor combination of  $\mathbf{Z}^\nu = \{1,10,13\}$  has better performance for values 1 and 2 than those from the combination  $\mathbf{Z}^\nu = \{1,3,10\}$ . Conversely the latter performs better for value 3. Using two-tailed hypothesis tests, setting  $H_0$  and  $\alpha$  as before, and taking into account the different data set sizes, the  $P$ -values were found to be 0 for all three values. This means the performance differences are significant. Now, Tables 5 and 6 show that the differences between the prediction results of values 2 and 3 are small compared with those of value 1. Therefore, the sensor combination of  $\mathbf{Z}^\nu = \{1,10,13\}$  gives a better overall performance than that of  $\mathbf{Z}^\nu = \{1,3,10\}$ .

Table 5: Inference results for sensor combination of nodes 1, 10 and 13.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	2618	<b>100</b>	0	0
2	4460	0	<b>99.57</b>	0.43
3	5281	0	1.61	<b>98.39</b>

Table 6: Inference results for sensor combination of nodes 1, 3 and 10.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	1514	<b>93.39</b>	6.61	0
2	2513	0	<b>98.81</b>	1.19
3	2634	0	0.53	<b>99.47</b>

The prediction results from the two sensor layouts deduced for four sensors are shown in Tables 7 and 8. These results show similar differences as those for three sensors. Similarly, two-tailed hypothesis tests show these differences are statistically significant. Thus, the sensor layout obtained by the MI criterion (nodes 2, 5, 10 and 13) perform worse overall than the layout obtained from the other three criteria (nodes 1, 8, 10, 13).



Table 7: Inference results for sensor combination of node 1, 8, 10 and 13.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	921	<b>100</b>	0	0
2	1482	0	<b>99.39</b>	0.61
3	1437	0	1.46	<b>98.54</b>

Table 8: Inference results for sensor combination of node 2, 5, 10 and 13.

Ground truth		Prediction (%)		
Value	n	1	2	3
1	1501	<b>85.54</b>	14.46	0
2	1368	0	<b>99.85</b>	0.15
3	2437	0	4.76	<b>95.24</b>

## 6. Discussion and Conclusions

This report describes the first stages of work aimed at developing a methodology for determination of optimal sensor layouts: it has considered methods for placing a limited number of sensor nodes in an environment of interest such that the cost of the placement is minimised while the value of the obtained information is maximised. The formalism developed allows for the incorporation of actual costs or cost functions if and when they are known. These include the costs associated with deploying sensors in the environment  $K$ , and the costs of carrying out (or not carrying out) specific actions as a result of the sensed information. The approach presented here of defining the optimal expected cost of a sensor in terms of the information it provides has been shown to predict optimal sensor layouts accurately.

Specifically, this initial work has focused on direct measurements, that is, sensor layouts where sensors are placed in only a subset of possible locations, leaving the rest of the space without sensors. Four criteria were compared: the criterion developed here that in this case reduces to maximum entropy of sensor measurements; minimum aggregated residue entropy (maximum reward); maximum mutual information (MI) between the sensors and the rest of the space; and maximum information coverage (IC).

Verification was carried out using data from an existing wireless sensor network. An environment model was learnt as a Bayesian Network. Each criterion was applied independently producing in general different optimal sensor layouts. It was found that for three or more sensors deployed in a layout, all four criteria placed some sensors on the edges and some near the middle of the area. Furthermore, the maximum MI criterion was observed to be the only one that gives a different layout. Each layout was used to predict sensor measurements in the rest of the space.

To verify the performance of the layouts, the sensor data not utilised in learning was used as the ground truth by comparing it with the predicted measurements. For all sensor combinations it was found that those learnt from using the entropy and IC criteria gave the best performance results.

It is interesting to note that the predictive ability of the two-sensor combination at nodes 1 and 13 (Table 1) is apparently very similar to that of the four-sensor

combination at nodes 1, 8, 10 and 13 (Table 7), i.e. the addition of two additional sensors at nodes 8 and 10 provided little tangible improvement in the results. This seems (intuitively) a little surprising in view of the structure of the network (Figure 6). The results outlined above focussed more on finding optimal layouts for a pre-determined number of sensors than on the equally important issue of determining the optimal number of sensors. This will be more explicitly addressed in future work.

To differentiate between the entropy and IC criteria, we may compare their computational complexities. Both criteria require  $N^{|V|}$  steps, where  $N$  is the size of  $\mathbf{X}$  and  $|V|$  is the number of sensors, thus giving us a value for every combination of  $|V|$  sensors. However, the IC criterion only requires pairwise entropies (entailing marginalisations over at most two sensor models), but the entropy criterion requires entropy computation for  $|V|$  sensors. Thus, these two criteria have the same computational complexity for two sensors, but the IC criterion has less complexity as the number of sensors increases. Therefore, arguably, when using discrete variables and direct measurements, the optimal sensor layout is best found using the information coverage criterion given both prediction accuracy and computational complexity.

However, care must be taken before drawing general conclusions from a single data set. As indicated in Section 5.1, further investigation is required of the conditions under which the IC criterion produces similar results to the entropy criterion suggested in this work.

It may be noted, perhaps trivially, that the more sensors that are used the more similar should be the optimal layouts deduced from the different criteria. For a full set of, in the present case, 16 sensors, there can be no difference in optimal layouts.

In addition to further investigation of the generality or otherwise of the results presented here, future work will include similar comparisons for situations using indirect measurements and extensions to different sensor networks and types of data.

A final comment concerns the future use of physical models to provide the prior knowledge needed to define an optimal sensor layout. Statistical models such as the Bayesian Network used in the example of Sections 4 and 5 make use of no prior knowledge about the environment other than that employed in deciding on the spatial separation of the nodes of the full sensor grid used for collecting the training data (Figure 4). This is appropriate in many cases, particularly for monitoring of natural structures, when prior knowledge of the environment is limited. However, we generally have quite detailed knowledge of engineered structures when they are new, though this knowledge may be incomplete in some important aspects. And of course our *a priori* knowledge will be reduced as the structure ages.

A challenging aspect of future work will be developing a methodology for incorporating the knowledge we have, which can be expressed as a physical model of the structure, with a method for learning where there are knowledge gaps due to inadvertent manufacturing variations and effects of ageing, into the formalism for identifying optimal sensor layouts. That is, we want to develop a methodology that incorporates appropriate aspects of both physical and sensor-based models into the formalism presented in this report. The approach to development of a hybrid model outlined by Cole *et al.* (2008) for corrosion monitoring is a promising direction for the future.

## References

- Cole, I.S., Corrigan, P.A., Ganther, W., Ho, T., Lewis, C.J., Muster, T.H., Paterson, D., Price, D.C., Scott, D.A., Followell, D., Galea, S., Hinton, B. (2008) "Development of a sensor-based learning approach to prognostics in intelligent vehicle health monitoring". In: International Conference on Prognostics and Health Management, Denver, CO, USA.
- Cowell, R.G., Dawid, A.P., Lauritzen, S.L., Spiegelhalter, D.J. (1999) "Probabilistic Networks and Expert Systems." Springer-Verlag, New York.
- Crutchfield, J.P. (1990) "Information and its metric". In: "Nonlinear Structures in Physical Systems. Pattern Formation, Chaos and Waves." pp.119-130. (eds. Lam, L., Morris, H.C.), Springer-Verlag.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977) "Maximum likelihood from incomplete data via the EM algorithm", J. Royal Statistical Soc. B, **39**, 1-39.
- Fawcett, T. (2003) "ROC graphs: notes and practical considerations for data mining researchers." Technical Report HPL-2003-4, HP Laboratories.
- González-Banos, H., Latombe, J.C. (2001) "A randomized art-gallery algorithm for sensor placement." In: SCG 01: Proceedings of the seventeenth annual symposium on computational geometry, New York, NY, USA. pp. 232-40.
- Guestrin, C., Krause, A., Singh, A. (2005) "Near-optimal sensor placements in Gaussian processes." In: Proceedings of 22<sup>nd</sup> International Conference on Machine Learning. pp. 265-72.
- Krause, A., Guestrin, C. (2005) "Near-optimal nonmyopic value of information in graphical models". In: Conference on Uncertainty in Artificial Intelligence.
- Krause, A., Guestrin, C. (2005a) "Optimal nonmyopic value of information in graphical models: efficient algorithms and theoretical limits". In: Proceedings of 19<sup>th</sup> International Joint Conference on Artificial Intelligence pp.1339-1345.
- Krause, A., Singh, A., Guestrin, C. (2008) "Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies" J. Machine Learning Res. **9**, 235-84.
- Lee, B.C., Staszewski, W.J. (2007) "Sensor location studies for damage detection with Lamb waves" Smart Mater. Struct. **16**, 399-408.
- MacKay, D. (2003) "Information Theory, Inference, and Learning Algorithms". Cambridge University Press.
- Myung, I.J. (2003) "Tutorial on maximum likelihood estimation", J. Math. Psychol. **47**, 90-100.
- Olsson, L., Nehaniv, C.L., Polani, D. (2004) "Information trade-offs and the evolution of sensory layout." In: Proceedings of the Ninth International Conference on Artificial Life, 119-124.
- Pearl, J. (1988) "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference." Morgan Kauffman.
- Ramakrishnan, N., Bailey-Kellogg, C., Tadepalli, S., Pandey, V.N. (2005) "Gaussian processes for active data mining of spatial aggregates." In: Proceedings of the SIAM International Conference on Data Mining.

- Shannon, C. (1948) "A mathematical theory of communication." Bell System Technical Journal **27**, 623-656.
- Staszewski, W.J., Worden, K., Wardle, R., Tomlinson, G.R. (2000) "Fail-safe sensor distributions for impact detection in composite materials" Smart Mater. Struct. **9**, 298-303.
- Trendafilova, I., Heylen, W., Van Brussel, H. (2001) "Measurement point selection in damage detection using the mutual information concept" Smart Mater. Struct. **10**, 528-33.
- Wang, X.R., Lizier, J.T., Obst, O., Prokopenko, M., Wang, P. (2008) "Spatiotemporal anomaly detection in gas monitoring sensor networks." In: Wireless Sensor Networks (ed. R. Verdone) Volume 4913, pp. 90-105. Springer, Berlin.
- Yeung, R.W. (1991) "A new outlook of Shannon's information measures" IEEE Trans Information Theory **37** (3), 466-74.

## Appendix 1

### Sentient Structures

White Paper for US Air Force, Asian Office of Aerospace R&D  
from  
**Commonwealth Scientific & Industrial Research Organisation**  
(CSIRO Australia)



### **Summary**

Engineering structures of the future, whether they are vehicles, buildings, infrastructure, networks, etc., or heterogeneous groups of such structures, will be required to be perceptive and responsive. Such structures will be referred to as sentient structures. This White Paper discusses the development of techniques and technologies by which sentient structures can formulate perceptions, awareness and intelligent responses to events and environments that may cause, or have caused, damage to the structure.

The approach to be pursued is to distribute sensing, active response and computational capabilities throughout the structure to form a complex multi-agent network, and to develop diagnostic, prognostic and decision-making functions entirely by self-organization<sup>1</sup> of the complex system, with no central control. Such an approach will yield robust, adaptive and scaleable systems.

A key focus in the early stages of the work will be the development of information-theoretic techniques for determining optimal sensor densities and layouts, suitable for a fully distributed environment, for specific damage formation and propagation processes in real materials. A hardware demonstrator has been developed to enable simulation on a real distributed system, forming a bridge between computer-based simulation and application-specific prototypes.

### **Introduction**

Structural health monitoring and management (SHM) is a new approach to assuring the fitness for purpose of critical structures. SHM employs sensors built into a structure to continuously monitor its state. It will initially reduce the need for, and may ultimately replace, the current regime of periodic non-destructive inspection and evaluation (NDE).

Current SHM systems, which are essentially experimental, are relatively narrowly focussed on particular damage “hot spots” in structures such as aircraft. If SHM systems are to be more broadly based, key requirements will be an ability to process data from a large number of sensors in different parts of the structure, and to continue performing effectively in the presence of damage.

The approach we have adopted to satisfy these requirements is a distributed multi-agent system, in which semi-autonomous local agents control a suite of sensors and process their data to obtain information about the state of the structure in its local region. These local agents communicate with their neighbours, with the objective of the system as a whole forming a diagnosis of the damage, and ultimately a response to it, by the process of self-organisation. We have recently demonstrated for the first time an example of such a self-organising SHM system. This is outlined in a recent publication of ours (Hoschke et.al., 2007), which is attached since it is not yet available.

---

<sup>1</sup> “Self-organization is a process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system’s components are executed using only local information, without reference to the global pattern”, Camazine *et al.* *Self-Organization in Biological Systems*, Princeton University Press (2001).

The long-term aim of this approach is the development of sentient structures. Sentient structures will be able to sense (or “feel”) damage as it occurs, to evaluate the nature and severity of damage, to infer its cause (diagnosis), and to make a prediction (prognosis) of damage development and its effect on the performance of the structure in the future. They will have the capability to make decisions for remedial actions, ultimately including self-repair. They will be aware of environmental or operational conditions that may cause damage. Sentient structures will fail only in extreme circumstances, because many of the common causes of failure will be detected and corrected at an early stage.

Sentient structures will have a major impact in many areas, including transportation (e.g. space vehicles, aircraft, motor vehicles), heavy machinery (such as mining, manufacturing and processing equipment), buildings and infrastructure (e.g. dams, bridges, pipelines and networks), and the protection of critical infrastructure. Sentient structures will be safer, will greatly reduce maintenance costs, and, most significantly, will allow the use of more efficient structural designs.

The major benefits of the SHM approach will follow the development of materials that have inherent sensory, and eventually self-healing, capabilities, integrated communications and processing elements, and robust, intelligent systems capable of processing a vast amount of data, learning, adapting, and formulating intelligent responses to the threat or occurrence of damage.

One of our immediate objectives is the development of a rigorous approach to the design of optimal sensor layouts. Information and information flows are the key ingredients in the design of distributed SHM systems, so information-theoretic and probabilistic inference techniques will be applied to the problem of designing efficient and effective sensing systems. Attention will be focussed on the development of revolutionary capabilities for diagnosis and prognosis of the structure within a fully distributed sensing and computational network, by designing the self-organized response of the network to specific damage scenarios.

Much international research in SHM is either aimed at incremental developments for the deployment of near-term technology, or is focussed on specific aspects of the problem (sensor development, in particular). Incremental developments, and their near-term applications, are very important for a variety of reasons (including accumulation of domain knowledge, and gaining of industry acceptance) but they are not the focus of this project.

### **CSIRO Capability for this Research**

CSIRO is an Australian Government research and development organisation, with broad interests and capabilities across many areas of science and technology. It employs some 5000 technical staff in 20 research divisions, providing a powerful ability to form strong multi-disciplinary teams to tackle significant problems. Further details about CSIRO can be found at <http://www.csiro.au>.

CSIRO has a number of research activities that are either directly involved with or closely relevant to the work proposed here. A collaboration between groups in two Divisions, CSIRO Industrial Physics (CIP) and the CSIRO ICT Centre (CICTC), has been working for some four years on the development and demonstration of concepts for the intelligent systems aspects of structural health management, in a project partially supported by NASA (Langley Research Center) and The Boeing Company. The multi-disciplinary project team draws on existing expertise in sensing, NDE, signal processing, telecommunications and intelligent systems. A recent paper (Hoschke, et.al., 2007) that outlines the approach adopted, current progress and some recent results is attached. A hardware test-bed/demonstrator that contains 192 autonomous sensing agents that form a complex multi-agent system has been developed as part of this work.

CSIRO also has research activities and capabilities in the design and development of functional composite materials (which includes contract work for Boeing, with whom CSIRO has an active R&D partnership), sensors and sensing, nanoscience, damage in metals and

composites, intelligent textiles, processing of very large data sets, etc. The aim is to bring some of these capabilities, along with those of external partners, into the present project at a later stage (see below).

## **Major Objectives and Approach**

*Ad hoc* networks and pervasive computing are active areas of research worldwide, but the detailed diagnostic and prognostic issues involved in SHM have received little research attention. There is as yet no general successful approach to the engineering of complex systems such as sentient structures to produce desired self-organized outcomes.

In order to provide a revolutionary rather than incremental advancement toward sentient structures, we propose to develop the following concepts implemented in a hardware concept demonstrator:

- 1) a multi-cellular sensor and communication network, including optimal sensor layouts, flexible communication and coordination mechanisms, and self-maintenance capabilities, based on a novel evolutionary design methodology;
- 2) a self-organizing response system, utilizing distributed sensor data from the multi-cellular network with results from internal damage models (diagnostic, prognostic, etc.), and supporting decentralised decision-fusion within the network;
- 3) novel verification and validation techniques for decentralised distributed systems, using information-theoretic metrics to quantitatively measure design outcomes and performance of self-organizing systems with non-deterministic emergent behaviour.

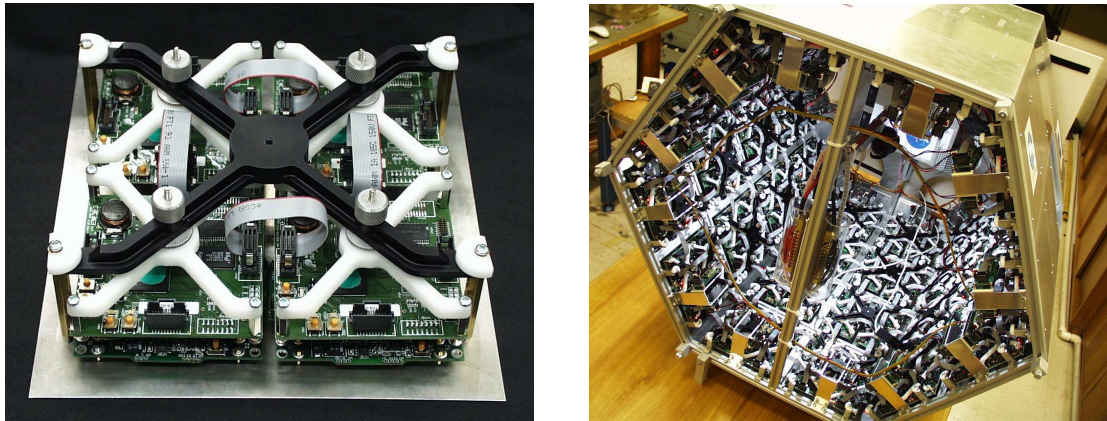
### 1. Evolutionary design of a distributed sensing network

The purpose of the proposed multi-cellular network is to provide a flexible, modular, reconfigurable skin for a sentient structure, embedding cells with multiple sensing modalities, collectively capable of a wide range of self-assessment functions. The multi-cellular network will deal with both simultaneous, real-time events (e.g. impacts), and long-horizon transients (e.g. material degradation such as corrosion or fatigue).

The transition from conventional “hot spot” monitoring, which uses relatively few sensors and treats damage detection as a separate task from data analysis and prognosis, to SHM that will employ very large numbers of diverse sensors integrated into the material microstructure, will necessitate handling of massive amounts of data. These systems have to be designed comprehensively, aiming at optimal sensor densities and layouts, adequate information transfer between the system’s components, and reliable inference for diagnostics, prognostics and response. Given the requirements of robustness, adaptability and scalability, these tasks cannot be achieved with traditional engineering methods that result in segmented, brittle designs incapable of adapting to new situations. By contrast, biological systems are not built out of separately designed parts attached together at a later stage – they evolve symbiotically. Each component is reliant on other components and co-evolves to work even more closely with the whole. The result is a dynamic system where components can be reused for other purposes and take on multiple roles, increasing robustness observed on multiple levels: from a cell to an ant colony to social systems (Miller et.al., 2000).

In order to approach the required levels of robustness, adaptability and scalability in solving the SHM problems, we propose a biologically-inspired multi-cellular sensor and communication network, with self-monitoring and self-diagnosing capabilities, aiming at self-organizing response. Data will be processed locally, and only information relevant to other regions of the structure will be communicated.

The main network component is an autonomous cell: a multi-layered hardware module, including layers for external protection, embedded sensors, electronic data acquisition, software for communications, power distribution and agent behaviors. It has a limited number of communication/power connections to neighboring cells covering a given surface. Cells in the existing demonstrator, referred to above, are shown in Figure 1.



*Figure 1: The current CSIRO multi-cellular network experimental test-bed/demonstrator. The image on the left shows four cells attached to an aluminium “skin” panel, while that on the right shows the assembled demonstrator that consists of 192 cells.*

The inter-connected cells create a network without centralized controllers. The cells can be manufactured or retrieved independently without any knowledge of the network topology or state, and can be added to the network anytime, anywhere, resulting in highly scalable SHM systems. The network is able to continue functioning when some individual cells are destroyed or malfunction. This is achieved by localized algorithms, using only local behaviors and communications. Without centralized controllers, cells deal with regional failures, resulting in highly robust reconfigurable SHM systems.

In designing self-organizing systems we may draw to some degree from traditional engineering top-down decomposition design methods, classical AI planning and reasoning techniques, bottom-up emergent behaviour engineering (such as reaction-diffusion, amorphous computing, graph automata), but essentially require a methodology for a co-evolution of multiple agents fitting selection criteria collectively – as a multi-agent system. Typically, evolutionary design may employ genetic algorithms in evolving optimal strategies that satisfy given fitness functions, by exploring large and sophisticated search-space landscapes (Miller et.al., 2000). Nevertheless, we may approach evolutionary design in two ways: via task-specific objectives or via generic intrinsic selection criteria (Prokopenko et.al., 2006a). The latter method – information-driven evolutionary design – essentially focuses on information transfer within specific channels. An example of an information-theoretic selection pressure is the acquisition of information from the environment: there is some evidence that pushing the information flow to the information-theoretic limit (i.e., maximisation of information transfer) can give rise to intricate behaviour, induce a necessary structure in the system, and ultimately be responsible for adaptively reshaping the system (Kluybin et.al., 2004).

In a distributed scenario the information-driven evolutionary design question becomes: what are the co-evolving sensors, actuators, memory states, and behaviours which maximize the information-transfer in a given dynamic environment. In particular, we intend to evolve

- optimal sensor layouts;
- optimal inference from observation to diagnosis;
- optimal inference from diagnosis to prognosis; and
- coordinated perception-action loops.

## 2. Self-organising response

The development of a self-organising response (e.g. self-repair or autonomous maintenance scheduling) in a distributed multi-agent network is the immediate focus of this project. The nature of the response must be determined by the damage diagnosis and prognosis: what is the



nature of the past events, how does the damage affect the functionality of the structure now and in the future, and when might it become critical?

The purpose of our method is not to provide an exact description of damage processes and possible reactions, but to provide a sufficiently realistic model to use for the development of a response strategy. In this context the inherent uncertainty of diagnostics and prognostics resulting from uncertainties in the knowledge of the past and the future operating conditions is recognised.

Single cells may make fast and automatic responses to critical emergencies, while collections of cells may solve more complex hierarchical tasks, for example:

- a) self-calibrate, discriminate among component and sensor failures;
- b) form a dynamic network, characterizing the nature of possible damage and inferring a self-organizing diagnosis and prognosis;
- c) self-schedule secondary inspections, maintenance or corrective actions based on information from the network, while issuing warnings;
- d) direct repair or recovery resources, human or robotic, to the repair site.

The establishment of an adequate set of the information-theoretic criteria will support a set of design guidelines for self-organizing sentient structures, applicable to a large-scale model system to be developed during the next stage of the project.

### 3. Verification and validation metrics

Condition-Based Maintenance (CBM) has become popular for complicated multi-component systems due to its cost and reliability advantages over traditional scheduled maintenance programs: for example, advanced reasoning schemes for collecting diagnostic/prognostic information and reducing false alerts are being developed in

at Pennsylvania State University (ARL-PSU). However, according to a NASA Jet Propulsion Laboratory report on Prognostics Methodology for Complex Systems (Gulati and Mackey, 2003), CBM is frequently difficult to apply to complex systems exhibiting emergent behaviour and facing highly stochastic environmental effects. A scalable solution capable of providing a substantial look-ahead capability is required. The JPL solution involves an automatic method to schedule maintenance and repair, targeting the two fundamental problems in autonomic logistics: (1) unambiguous detection of deterioration or impending loss of function and (2) determination of the time remaining to perform maintenance or other corrective action based upon information from the system (Gulati and Mackey, 2003). The solution based on the JPL work, nevertheless, does not account for self-organization and is not directly applicable to distributed multi-agent networks.

Most engineering systems being designed today are very large, distributed, decentralised, and complex. A distinguishing feature of complex systems is the emergence of system-level behaviour out of the interactions among local nodes. Traditional multi-component systems do not exhibit *self-organization*. Instead, they rely on fixed multiple links among the components in order to efficiently control the system, having fairly predictable and often pre-optimised properties, at the expense of being less scaleable and less robust. Consequently, the traditional verification methodology developed so far has very limited applicability with respect to complex systems: it does not capture self-organization and cannot fully measure resilience, fault-tolerance and recovery.

A new, promising approach to verification of complex systems suggests the use of information-theoretic metrics in measuring self-organization. Over the last 3 years we have investigated various information-theoretic measures (such as Shannon entropy of certain frequency distributions), targeting response time as well as spatial connectivity, temporal persistence and size of self-organising patterns (Prokopenko et.al. 2006b, 2005a, b, c, Hoschke et.al., 2007). These metrics may form a core of new verification methodology applicable to non-deterministic emergent behaviour, and measuring reliability and resilience of complex distributed and decentralised systems. There are very few research groups world-

wide that investigate the applicability of information-theoretic metrics to verification of scalable networked systems: on one hand, the effort is focused on theoretical aspects of the emergent behaviour, while on the other hand, “best practice” verification tools do not capture emergent behaviour. The proposed intersection fills an important and well-defined niche and will have a major impact on the development of sentient structures.

### **Benefits/payoff of the technology**

Both the initial focus of self-organized diagnosis and prognosis, and the broader objective of sentient structures, can clearly find important applications to many areas of defence assets and operations. The initial target application for this technology is the structural health management of vehicles and infrastructure, and in this application area it would lead to, *inter alia*, maintenance cost savings, improved structural reliability and efficiency, and enhanced personnel safety. However, the technology has much broader potential applications in areas in which robust, distributed situational diagnosis, prognosis and decision-making are required: examples include emergency response advisory systems, physical security of structures, networks, etc. All of these are of considerable relevance to Defence, and offer the possibility of revolutionary Defence capabilities. It should be recognised, of course, that significant integration efforts will be required to produce practical systems (see, e.g., Prosser et.al., 2004).

### **General Outline of Program, Costs and Duration**

This proposal is for an initial 12-month “seed” project, which aims to develop a formalism for designing optimum sensor distributions and layouts to enable a system or a local agent to efficiently acquire the information it needs for an appropriate response to be developed. This work will be based on information-theoretic principles, building on and extending the work of Polani’s group (e.g. Olsson et.al., 2004) by incorporating Bayesian inference of the damage state in a decentralised environment. Different approaches to optimising the design will be investigated. This is work that should ultimately be broadly applicable to distributed networks in a range of applications.

This work could be carried out by a combination of our existing research team staff, one or two PhD students or part of the time of a Postdoctoral Fellow. The make-up of the team will depend on a number of factors, and will be decided prior to the work commencing. The cost for any of the staffing options is expected to be approximately US\$50k for a 12-month project.

### **Expertise of the principal investigators**

It is proposed that the effort is led by Dr Mikhail Prokopenko (CSIRO ICT Centre) and Dr Don Price (CSIRO Industrial Physics).

Dr Mikhail Prokopenko has a strong international reputation in the areas of complex multi-agent systems and distributed intelligence (over 70 publications and patents). He received a PhD in Computer Science (Macquarie University, 2002, Australia), MA in Economics (University of Missouri-Columbia, 1994, USA), and MSc in Applied Mathematics (Azerbaijan Institute of Petroleum & Chemistry, 1988, USSR). Since joining CSIRO he has led a number of R&D projects, including a CSIRO Complex Systems Science Emerging Science Project on Directed Self-Assembly in Multi-Agent Networks (January 2003 – June 2004). In June 2002, Dr Prokopenko received the Japanese Society for Artificial Intelligence award for scientific contribution to the RoboCup Simulation League, for his work on entropy of joint beliefs as a measure of multi-agent coordination potential. Dr Prokopenko has worked on a number of international Program and Organising Committees; was a keynote speaker at 6th International Workshop on Agent-Based Simulation (2005); co-chaired sessions on Evolutionary and Self-Organizing Sensors, Actuators and Processing Hardware (International Conferences on Knowledge-Based Intelligent Information & Engineering Systems). He is an adjunct Associate Professor at the School of Computer Science and Engineering, the University of New South Wales (see also <http://www.ict.csiro.au/staff/Mikhail.Prokopenko/>).

Dr Don Price is currently Research Group Leader in CSIRO Industrial Physics, and Project Leader of the CSIRO-NASA Ageless Aerospace Vehicle Project, a collaboration that is developing and demonstrating systems concepts and techniques for advanced structural health management systems. He has an extensive research background in condensed matter physics, industrial applications of ultrasound (mainly in non-destructive evaluation of aerospace materials and structures) and more recently in complex self-organizing systems. This work has involved substantial collaborations with, *inter alia*, AEA Technology (Harwell, UK), Metal Manufactures (Port Kembla, NSW), the Boeing Company (Seattle and St.Louis, USA) and NASA (Langley Research Center, Dryden Flight Center, USA). His work over the past five years in structural health management has resulted in 33 publications, 2 invitations for keynote presentations at international workshops, and 2 invited book chapters. Dr Price was invited to join the NASA Engineering & Safety Center NDE Super Problem Resolution Team, which is investigating NDE issues for return to flight of the Space Shuttle, and the ISS.

## References

- APL-PSU. NASA Condition-Based Maintenance (NASA CBM). Applied Research Laboratory, The Pennsylvania State University, 2006.  
[http://www.arl.psu.edu/capabilities/mm\\_soa\\_research.html](http://www.arl.psu.edu/capabilities/mm_soa_research.html)
- Gulati, S. and Mackey, R. (2003). *Prognostics Methodology for Complex Systems. Automatic method to detect and react to complex degradation and incipient faults*. NASA Jet Propulsion Laboratory, Pasadena, California, NPO-20831.  
<http://www.nasatech.com/Briefs/July03/NPO20831.html>
- Hoschke, N., Lewis, C. J., Price, D. C., Scott, D. A., Gerasimov, V., Wang, P. (2007). A Self-organizing Sensing System for Structural Health Monitoring of Aerospace Vehicles. In Prokopenko, M., editor, *Advances in Applied Self-organizing Systems*, Springer-Verlag, London, UK, 2007 (accepted).
- Klyubin, A. S., Polani, D., Nehaniv, C. L. (2004). Organization of the Information Flow in the Perception-Action Loop of Evolved Agents. In *2004 NASA/DoD Conference on Evolvable Hardware, IEEE Computer Society, Seattle, Washington, June 24-26, 2004*, 177 – 180.  
<http://homepages.feis.herts.ac.uk/~comqdp1/publications/files/eh2004.ps.gz>
- Mathews, G. M., Durrant-Whyte, H. F., Prokopenko, M. (2006). Scalable Decentralised Decision Making and Optimisation in Heterogeneous Teams, IEEE International Conference on Multi-Sensor Fusion and Integration for Intelligent Systems (MFI-2006), Heidelberg, Germany, 383-388, September, 2006.  
<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/Mathews2006-MFI-ScalableDecisionMaking.pdf>
- Miller, J. F., Job, D., and Vassilev, V. K. (2000). Principles in the Evolutionary Design of Digital Circuits – Part I, *Journal of Genetic Programming and Evolvable Machines*, 1(1):8-35, 2000.  
<http://www.elec.york.ac.uk/intsys/users/jfm7/jgpem00a.pdf>
- Olsson, L., Nehaniv, C. L., Polani, D., Information Trade-offs and the Evolution of Sensory Layouts, In *Proceedings of the 9<sup>th</sup> International Conference on Artificial Life (ALife IX)*, Boston, Sept. 2004.  
[http://homepages.feis.herts.ac.uk/~comqdp1/publications/files/alife9\\_2004\\_final\\_olsson.pdf](http://homepages.feis.herts.ac.uk/~comqdp1/publications/files/alife9_2004_final_olsson.pdf)
- Prokopenko, M., Gerasimov, V., and Taney, I. (2006a). Evolving spatiotemporal coordination in a modular robotic system. In Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J. C. T., Marocco, D., Meyer, J.-A., Miglino, O., and Parisi, D., editors, *From Animals to Animats 9: 9th International Conference on the Simulation of Adaptive Behavior (SAB 2006)*, Rome, Italy, September 25-29 2006, volume 4095 of Lecture notes in computer science, 558–569.  
<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/snakebot-corrected.pdf>

Prokopenko, M., Poulton, G. T., Price, D. C., Wang, P., Valencia, P., Hoschke, N., Farmer, A. J., Hedley, M., Lewis, C., and Scott, D. A. (2006b). Self-organising impact sensing networks in robust aerospace vehicles. In Fulcher, J., editor, *Advances in Applied Artificial Intelligence*, 186–233. Idea Group, Hershey, PA, USA.

<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/AdvancesAAI-chapter.pdf>

Prokopenko, M., Wang, P., Foreman, M., Valencia, P., Price, D. C., and Poulton, G. T. (2005a). On connectivity of reconfigurable impact networks in ageless aerospace vehicles. *Journal of Robotics and Autonomous Systems*, 53(1):36–58.

<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/Connectivity-JRAS-published.pdf>

Prokopenko, M., Wang, P., and Price, D. C. (2005b). Complexity metrics for self-monitoring impact sensing networks. In Lohn, J., Gwaltney, D., Hornby, G., Zebulum, R., Keymeulen, D., and Stoica, A., editors, *Proceedings of 2005 NASA/DoD Conference on Evolvable Hardware (EH-05)*, Washington D.C., USA, 29 June – 1 July 2005, 239–246, IEEE Computer Society, Los Alamitos, CA, USA.

<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/e4-corrected.pdf>

Prokopenko, M., Wang, P., Price, D. C., Valencia, P., Foreman, M., and Farmer, A. J. (2005c). Self-organizing hierarchies in sensor and communication networks. *Artificial Life*, Special Issue on Dynamic Hierarchies, 11(4):407–426.

<http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/Hierarchies-Alife-proofs.pdf>

Prosser, W. H., Allison, S. G., Woodard, S. E., Wincheski, R. A., Cooper, E. G., Price, D. C., Hedley, M., Prokopenko, M., Scott, D. A., Tessler, A., Spangler, J. L. (2004). Structural Health Management for Future Aerospace Vehicles. In *The 2nd Australasian Workshop on Structural Health Monitoring (2AWSHM)*, Melbourne, Australia, December 2004, 1-15.

[http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/Australasian\\_SHM\\_Prosser.doc](http://www.ict.csiro.au/staff/mikhail.prokopenko/Publications/Agents/Australasian_SHM_Prosser.doc)